# An Ensemble Data Assimilation System for POP

B.T. Nadiga[1] & W.R. Casper[2]

Los Alamos National Lab
[1]balu@lanl.gov; public.lanl.gov/balu [2]wcasper@lanl.gov

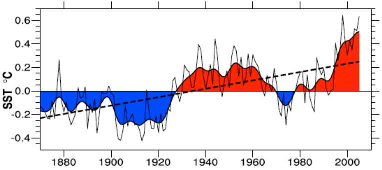
GODAE Oceanview 2011, Santa Cruz

(Thanks to Jeff Anderson, Nancy Collins, & Tim Hoar of NCAR)

# Outline

# Motivation

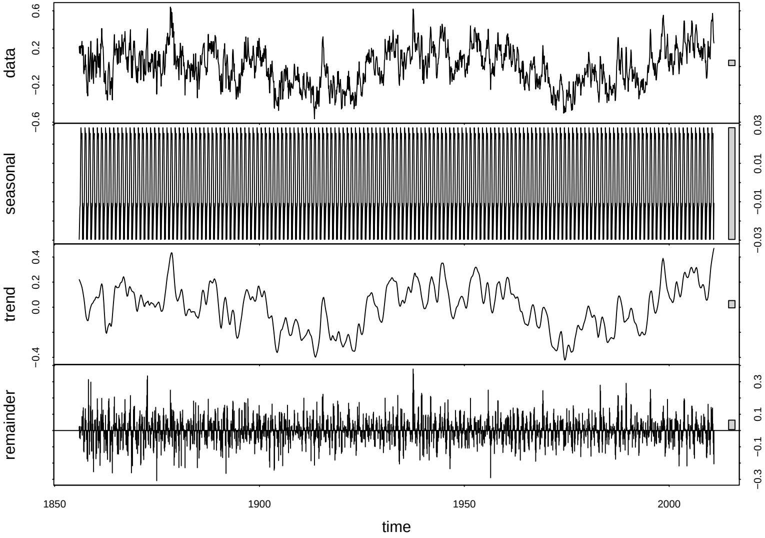
- Given its dynamical inertia, certain slow modes of global ocean circulation (e.g., AMO) are expected to be predictable on the interannual to decadal timescale



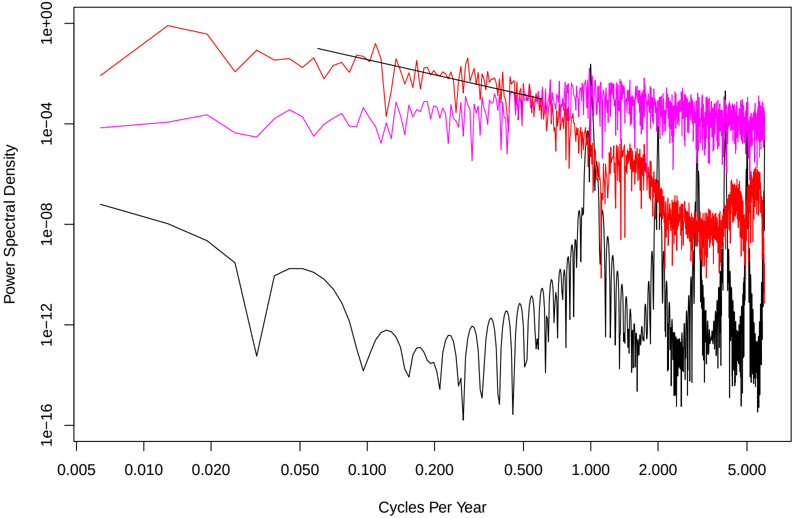

From Trenberth & Shea, 2006: Annual SST anomalies averaged over the North Atlantic (0 to 60 N, 0 to 80 W) for 1870–2005, relative to 1901–1970 (C)

- A pre-requisite to using the "extended predictability of slow modes" is a successful assimilation of data to estimate the state of the ocean including the phase and amplitude of the slow modes.
- Given recent improvements in methodology and other reasons, we have chosen to develop an ensemble DA system for POP

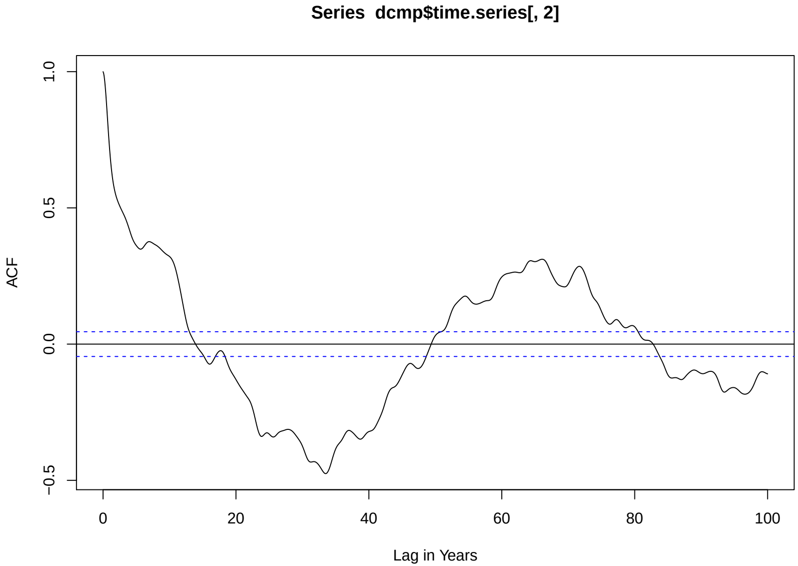
# Atlantic Multi-decadal Oscillation



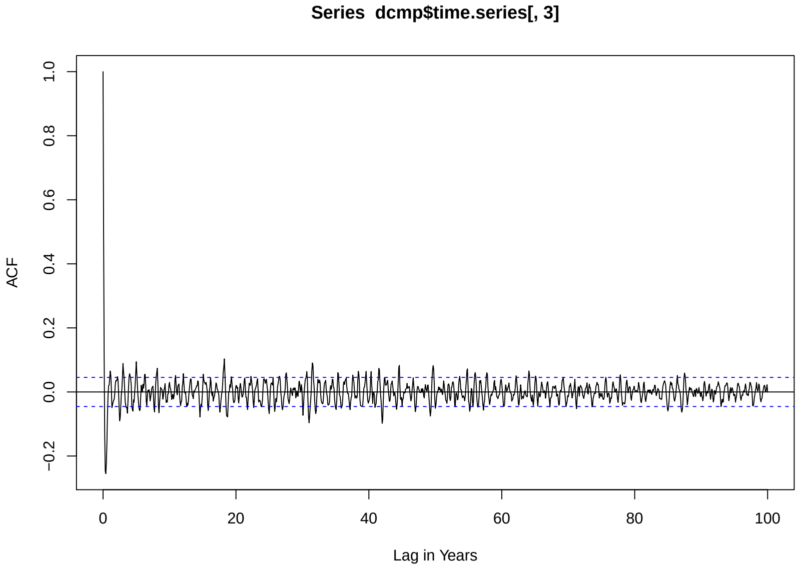
# Atlantic Multi-decadal Oscillation



# Atlantic Multi-decadal Oscillation



# Atlantic Multi-decadal Oscillation



# Assimilation Algorithm I

For ob. $y_{ob}$ with ob+representation error $\sigma_{ob}$,
and given prior ensemble $\mathbf{y} = h(\mathbf{x}_m)$ (bold: ensemble vector)
$m = 1, \ldots, state\_dim$. Use least squares:

- Compute posterior spread: $\frac{1}{\sigma_{po}^2} = \frac{1}{\sigma_{pr}^2} + \frac{1}{\sigma_{ob}^2}$
- Compute posterior ensemble mean: $\frac{\overline{y}_{po}}{\sigma_{po}^2} = \frac{\overline{y}_{pr}}{\sigma_{pr}^2} + \frac{y_{ob}}{\sigma_{ob}^2}$
- Compute ob. incr. for ensemble members (shift & compact): $\Delta\mathbf{y} = \overline{y}_{po} - \overline{y}_{pr} + \frac{\sigma_{po}}{\sigma_{pr}}\Delta\mathbf{y}_{pr}$
- Regress ob. incr. onto state variable incr. $\Delta\mathbf{x}_m = \beta(\mathbf{y}, \mathbf{x}_m)\Delta\mathbf{y}$ where $\beta(\mathbf{y}, \mathbf{x}_m) = cov(\mathbf{y}, \mathbf{x}_m)/\sigma_{pr}^2$
  - $\overline{\mathbf{x}}_{m,po} = \overline{\mathbf{x}}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\left(\overline{y}_{po} - \overline{y}_{pr}\right)$
  - $\Delta\mathbf{x}_{m,po} = \Delta\mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta\mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm I

For ob. $y_{ob}$ with ob+representation error $\sigma_{ob}$,
and given prior ensemble $\mathbf{y} = h(\mathbf{x}_m)$ (bold: ensemble vector)
$m = 1, \ldots, state\_dim$. Use least squares:

- Compute posterior spread: $\frac{1}{\sigma_{po}^2} = \frac{1}{\sigma_{pr}^2} + \frac{1}{\sigma_{ob}^2}$
- Compute posterior ensemble mean: $\frac{\overline{y}_{po}}{\sigma_{po}^2} = \frac{\overline{y}_{pr}}{\sigma_{pr}^2} + \frac{y_{ob}}{\sigma_{ob}^2}$
- Compute ob. incr. for ensemble members (shift & compact):
  $\Delta\mathbf{y} = \overline{y}_{po} - \overline{y}_{pr} + \frac{\sigma_{po}}{\sigma_{pr}}\Delta\mathbf{y}_{pr}$
- Regress ob. incr. onto state variable incr.
  $\Delta\mathbf{x}_m = \beta(\mathbf{y}, \mathbf{x}_m)\Delta\mathbf{y}$
  where $\beta(\mathbf{y}, \mathbf{x}_m) = cov(\mathbf{y}, \mathbf{x}_m)/\sigma_{pr}^2$
  - $\overline{\mathbf{x}}_{m,po} = \overline{\mathbf{x}}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\left(\overline{y}_{po} - \overline{y}_{pr}\right)$
  - $\Delta\mathbf{x}_{m,po} = \Delta\mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta\mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm I

For ob. $y_{ob}$ with ob+representation error $\sigma_{ob}$,
and given prior ensemble $\mathbf{y} = h(\mathbf{x}_m)$ (bold: ensemble vector)
$m = 1, \ldots, state\_dim$. Use least squares:

- Compute posterior spread: $\frac{1}{\sigma_{po}^2} = \frac{1}{\sigma_{pr}^2} + \frac{1}{\sigma_{ob}^2}$
- Compute posterior ensemble mean: $\frac{\overline{y}_{po}}{\sigma_{po}^2} = \frac{\overline{y}_{pr}}{\sigma_{pr}^2} + \frac{y_{ob}}{\sigma_{ob}^2}$
- Compute ob. incr. for ensemble members (shift & compact):
  $\Delta\mathbf{y} = \overline{y}_{po} - \overline{y}_{pr} + \frac{\sigma_{po}}{\sigma_{pr}} \Delta\mathbf{y}_{pr}$
- Regress ob. incr. onto state variable incr.
  $\Delta\mathbf{x}_m = \beta(\mathbf{y}, \mathbf{x}_m)\Delta\mathbf{y}$
  where $\beta(\mathbf{y}, \mathbf{x}_m) = cov(\mathbf{y}, \mathbf{x}_m)/\sigma_{pr}^2$
  - $\overline{\mathbf{x}}_{m,po} = \overline{\mathbf{x}}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\left(\overline{y}_{po} - \overline{y}_{pr}\right)$
  - $\Delta\mathbf{x}_{m,po} = \Delta\mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta\mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm I

For ob. $y_{ob}$ with ob+representation error $\sigma_{ob}$,
and given prior ensemble $\mathbf{y} = h(\mathbf{x}_m)$ (bold: ensemble vector)
$m = 1, \ldots, state\_dim$. Use least squares:

- Compute posterior spread: $\frac{1}{\sigma_{po}^2} = \frac{1}{\sigma_{pr}^2} + \frac{1}{\sigma_{ob}^2}$
- Compute posterior ensemble mean: $\frac{\overline{y}_{po}}{\sigma_{po}^2} = \frac{\overline{y}_{pr}}{\sigma_{pr}^2} + \frac{y_{ob}}{\sigma_{ob}^2}$
- Compute ob. incr. for ensemble members (shift & compact):
  $\Delta\mathbf{y} = \overline{y}_{po} - \overline{y}_{pr} + \frac{\sigma_{po}}{\sigma_{pr}} \Delta\mathbf{y}_{pr}$
- Regress ob. incr. onto state variable incr.
  $\Delta\mathbf{x}_m = \beta(\mathbf{y}, \mathbf{x}_m)\Delta\mathbf{y}$
  where $\beta(\mathbf{y}, \mathbf{x}_m) = cov(\mathbf{y}, \mathbf{x}_m)/\sigma_{pr}^2$
  - $\overline{\mathbf{x}}_{m,po} = \overline{\mathbf{x}}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\left(\overline{y}_{po} - \overline{y}_{pr}\right)$
  - $\Delta\mathbf{x}_{m,po} = \Delta\mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta\mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm I

For ob. $y_{ob}$ with ob+representation error $\sigma_{ob}$,
and given prior ensemble $\mathbf{y} = h(\mathbf{x}_m)$ (bold: ensemble vector)
$m = 1, \ldots, state\_dim$. Use least squares:

- Compute posterior spread: $\frac{1}{\sigma_{po}^2} = \frac{1}{\sigma_{pr}^2} + \frac{1}{\sigma_{ob}^2}$
- Compute posterior ensemble mean: $\frac{\overline{y}_{po}}{\sigma_{po}^2} = \frac{\overline{y}_{pr}}{\sigma_{pr}^2} + \frac{y_{ob}}{\sigma_{ob}^2}$
- Compute ob. incr. for ensemble members (shift & compact):
  $\Delta\mathbf{y} = \overline{y}_{po} - \overline{y}_{pr} + \frac{\sigma_{po}}{\sigma_{pr}}\Delta\mathbf{y}_{pr}$
- Regress ob. incr. onto state variable incr.
  $\Delta\mathbf{x}_m = \beta(\mathbf{y}, \mathbf{x}_m)\Delta\mathbf{y}$
  where $\beta(\mathbf{y}, \mathbf{x}_m) = cov(\mathbf{y}, \mathbf{x}_m)/\sigma_{pr}^2$
  - $\overline{\mathbf{x}}_{m,po} = \overline{\mathbf{x}}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\left(\overline{y}_{po} - \overline{y}_{pr}\right)$
  - $\Delta\mathbf{x}_{m,po} = \Delta\mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta\mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm II

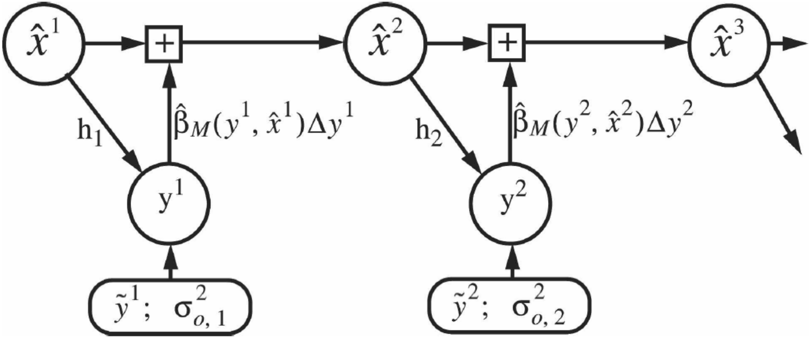



FIG. 1. A schematic depiction of the sequential filter algorithm. The forward observation operator for the first observation, $h_1$, is applied to the ensemble state vector, $\hat{\mathbf{x}}^1$, to produce a prior ensemble approximation, $\mathbf{y}^1$, of the observation. Observation increments, $\Delta\mathbf{y}^1$, are computed using the observation value, $\tilde{y}^1$, and error variance, $\sigma^2_{o,1}$, and regression is used to compute increments for the state, $\hat{\beta}_M(\mathbf{y}^1, \hat{\mathbf{x}}^1)\Delta\mathbf{y}^1$. The state is updated by adding the increments to produce $\hat{\mathbf{x}}^2$ and the process is repeated for each observation in turn.

(From Anderson & Collins, 2007)

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} = \mathbf{Z}^f \mathbf{C}\left(\Gamma + 1\right)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f \mathbf{C}\left(1 + \Gamma\right)^{-1/2} \mathbf{G}^{-1} \mathbf{F}^T \mathbf{Z}^f$ (20)

From the point of view of flow instabilities,
there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m) \frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta \mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} = \mathbf{Z}^f \mathbf{C} (\Gamma + 1)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f \mathbf{C} (1 + \Gamma)^{-1/2} \mathbf{G}^{-1} \mathbf{F}^T \mathbf{Z}^f$ (20)

From the point of view of flow instabilities,
there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m) \frac{\sigma_{po}}{\sigma_{pr}} h (\Delta \mathbf{x}_{m,pr})$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f\mathbf{T} = \mathbf{Z}^f\mathbf{C}\left(\Gamma + 1\right)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f\mathbf{C}\left(1 + \Gamma\right)^{-1/2}\mathbf{G}^{-1}\mathbf{F}^T\mathbf{Z}^f$ (20)

From the point of view of flow instabilities,
there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta \mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} = \mathbf{Z}^f \mathbf{C} \left(\Gamma + 1\right)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f \mathbf{C} \left(1 + \Gamma\right)^{-1/2} \mathbf{G}^{-1} \mathbf{F}^T \mathbf{Z}^f$ (20)

From the point of view of flow instabilities, there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m) \frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta \mathbf{x}_{m,pr}\right)$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f\mathbf{T} = \mathbf{Z}^f\mathbf{C}\,(\Gamma + 1)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f\mathbf{C}\,(1 + \Gamma)^{-1/2}\,\mathbf{G}^{-1}\mathbf{F}^T\mathbf{Z}^f$ (20)

From the point of view of flow instabilities,
there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m)\frac{\sigma_{po}}{\sigma_{pr}} h\,(\Delta \mathbf{x}_{m,pr})$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} = \mathbf{Z}^f \mathbf{C} (\Gamma + 1)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f \mathbf{C} (1 + \Gamma)^{-1/2} \mathbf{G}^{-1} \mathbf{F}^T \mathbf{Z}^f$ (20)

From the point of view of flow instabilities, there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
  - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m) \frac{\sigma_{po}}{\sigma_{pr}} h(\Delta \mathbf{x}_{m,pr})$

# Assimilation Algorithm III

Tippet et al., 2003 compare 3 deterministic SRFs

ETKF: $\mathbf{Z}^a = \mathbf{Z}^f \mathbf{T} = \mathbf{Z}^f \mathbf{C} \left(\Gamma + 1\right)^{-1/2}$ (16)

EAKF: $\mathbf{Z}^a = \mathbf{A}\mathbf{Z}^f = \mathbf{Z}^f \mathbf{C} \left(1 + \Gamma\right)^{-1/2} \mathbf{G}^{-1} \mathbf{F}^T \mathbf{Z}^f$ (20)

From the point of view of flow instabilities, there are important differences:

- Analysis perturbations in ETKF are linear combinations of (ens. no. of) forecast perturbations (each state vector component is similarly reconstituted/recombined.)
- Different state vector components of ensemble perturbation scale differently with EAKF (the adjustment matrix is of size state_dim x state_dim.)
    - $\Delta \mathbf{x}_{m,po} = \Delta \mathbf{x}_{m,pr} + \beta(\mathbf{y}, \mathbf{x}_m) \frac{\sigma_{po}}{\sigma_{pr}} h\left(\Delta \mathbf{x}_{m,pr}\right)$

# Spatio-Temporally Adaptive Inflation in DART

- Use $y_{ob}$, $\mathbf{y}$, and $\sigma_{ob}$ via Bayes theorem to improve $\lambda(x_m, t)$ [Anderson, 2009]
- Prior PDF for $\lambda(x_m)$ is multivariate normal. Sequentially update each inflation factor
- Assume $cov(\lambda(x_{m_1}), \lambda(x_{m_2})) = cov(x_{m_1}, x_{m_2})$
- No time evolution of inflation factor (Persistence)
- Assuming no bias, obtain posterior mean of $\lambda$ using approximations in Bayes formula; hold $\sigma_\lambda$ fixed $\mathrm{N}(\overline{\lambda}_{po}, \sigma_\lambda^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-D^2/2\theta^2\right) \mathrm{N}(\overline{\lambda}_{pr}, \sigma_\lambda^2)$ where $D = |\overline{y}_{pr} - y_{ob}|$ and $\theta = \sqrt{\overline{\lambda}_{po}\sigma_{pr}^2 + \sigma_{ob}^2}$

# Spatio-Temporally Adaptive Inflation in DART

- Use $y_{ob}$, $\mathbf{y}$, and $\sigma_{ob}$ via Bayes theorem to improve $\lambda(x_m, t)$ [Anderson, 2009]
- Prior PDF for $\lambda(x_m)$ is multivariate normal. Sequentially update each inflation factor
- Assume $cov(\lambda(x_{m_1}), \lambda(x_{m_2})) = cov(x_{m_1}, x_{m_2})$
- No time evolution of inflation factor (Persistence)
- Assuming no bias, obtain posterior mean of $\lambda$ using approximations in Bayes formula; hold $\sigma_\lambda$ fixed
  $\mathrm{N}(\overline{\lambda}_{po}, \sigma_\lambda^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-D^2/2\theta^2\right) \mathrm{N}(\overline{\lambda}_{pr}, \sigma_\lambda^2)$ where
  $D = |\overline{y}_{pr} - y_{ob}|$ and $\theta = \sqrt{\overline{\lambda}_{po}\sigma_{pr}^2 + \sigma_{ob}^2}$

# Spatio-Temporally Adaptive Inflation in DART

- Use $y_{ob}$, $\mathbf{y}$, and $\sigma_{ob}$ via Bayes theorem to improve $\lambda(x_m, t)$ [Anderson, 2009]
- Prior PDF for $\lambda(x_m)$ is multivariate normal. Sequentially update each inflation factor
- Assume $cov(\lambda(x_{m_1}), \lambda(x_{m_2})) = cov(x_{m_1}, x_{m_2})$
- No time evolution of inflation factor (Persistence)
- Assuming no bias, obtain posterior mean of $\lambda$ using approximations in Bayes formula; hold $\sigma_\lambda$ fixed $\mathrm{N}(\overline{\lambda}_{po}, \sigma_\lambda^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-D^2/2\theta^2\right) \mathrm{N}(\overline{\lambda}_{pr}, \sigma_\lambda^2)$ where $D = |\overline{y}_{pr} - y_{ob}|$ and $\theta = \sqrt{\overline{\lambda}_{po}\sigma_{pr}^2 + \sigma_{ob}^2}$

# Spatio-Temporally Adaptive Inflation in DART

- Use $y_{ob}$, $\mathbf{y}$, and $\sigma_{ob}$ via Bayes theorem to improve $\lambda(x_m, t)$ [Anderson, 2009]
- Prior PDF for $\lambda(x_m)$ is multivariate normal. Sequentially update each inflation factor
- Assume $cov(\lambda(x_{m_1}), \lambda(x_{m_2})) = cov(x_{m_1}, x_{m_2})$
- No time evolution of inflation factor (Persistence)
- Assuming no bias, obtain posterior mean of $\lambda$ using approximations in Bayes formula; hold $\sigma_\lambda$ fixed $\mathrm{N}(\overline{\lambda}_{po}, \sigma_\lambda^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-D^2/2\theta^2\right) \mathrm{N}(\overline{\lambda}_{pr}, \sigma_\lambda^2)$ where $D = |\overline{y}_{pr} - y_{ob}|$ and $\theta = \sqrt{\overline{\lambda}_{po}\sigma_{pr}^2 + \sigma_{ob}^2}$

# Spatio-Temporally Adaptive Inflation in DART

- Use $y_{ob}$, $\mathbf{y}$, and $\sigma_{ob}$ via Bayes theorem to improve $\lambda(x_m, t)$ [Anderson, 2009]
- Prior PDF for $\lambda(x_m)$ is multivariate normal. Sequentially update each inflation factor
- Assume $cov(\lambda(x_{m_1}), \lambda(x_{m_2})) = cov(x_{m_1}, x_{m_2})$
- No time evolution of inflation factor (Persistence)
- Assuming no bias, obtain posterior mean of $\lambda$ using approximations in Bayes formula; hold $\sigma_\lambda$ fixed $\mathrm{N}(\overline{\lambda}_{po}, \sigma_\lambda^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-D^2/2\theta^2\right) \mathrm{N}(\overline{\lambda}_{pr}, \sigma_\lambda^2)$ where $D = |\overline{y}_{pr} - y_{ob}|$ and $\theta = \sqrt{\overline{\lambda}_{po}\sigma_{pr}^2 + \sigma_{ob}^2}$

# Setup of the DA Experiment

- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Setup of the DA Experiment

- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Setup of the DA Experiment

- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Setup of the DA Experiment

- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Setup of the DA Experiment

- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Setup of the DA Experiment

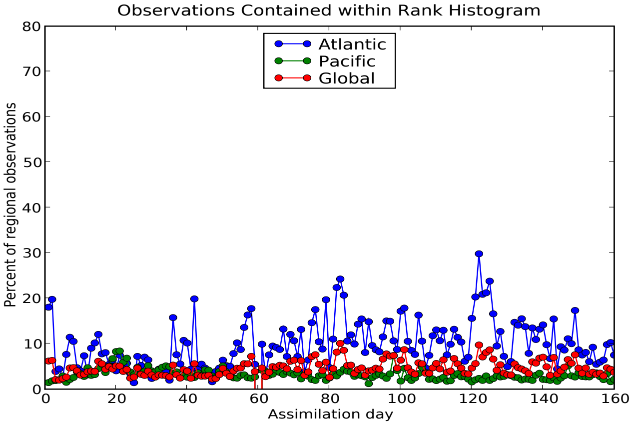
- LANL POP with Data Assimilation Research Testbed [Anderson et al., 2009]
- Corrected CORE (Coordinated Ocean-ice Reference Experiment) version 2 Interannual Forcing [Large and Yeager, 2009, Griffies et al., 2009]
- Weak salinity restoring; strong SST restoring under ice
- 20 member ensemble with spatiotemporally adaptive inflation [Anderson, 2009]; localization radius of 1100 km
- ICs from Jan 1 of different years of a control run with Normal Year Forcing
- World Ocean Database observations starting Jan 1, 1990
- An identical control ensemble run, but with no assimilation

# Analysis of 1990-91 DA Experiment I

DA system performs poorly with too few obs. ($< 10\%$) being contained in the ensemble

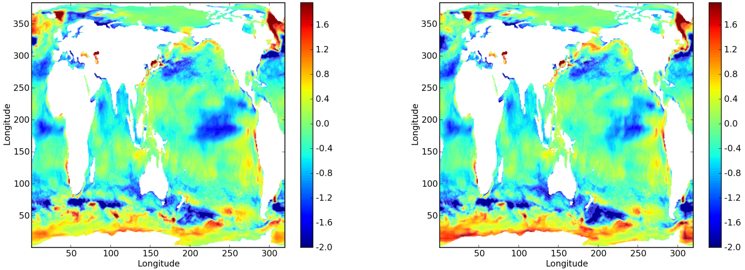
# Analysis of 1990-91 DA Experiment I

DA system performs poorly with too few obs. ($< 10\%$) being contained in the ensemble



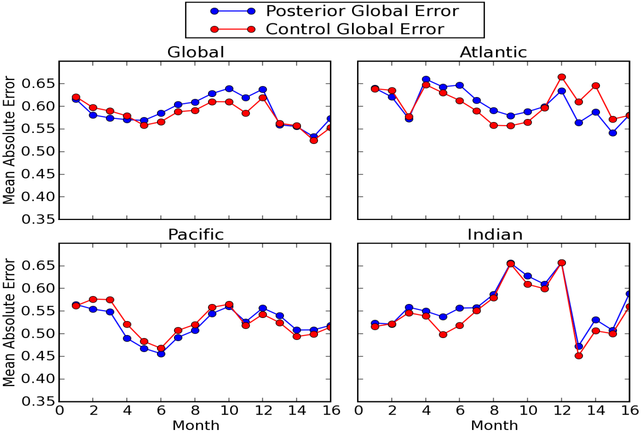

Insufficient ensemble spread in spite of sophisticated, spatio-temporally adaptive, statistical inflation [Anderson, 2009]

# Analysis of 1990-91 DA Experiment II




Time-avgd difference wrt NOAA OI SST v2 **Left**: Control Run; **Right**: Assimilation Run. Cold bias in tropics and midlatitudes and warm bias at high latitudes and upwelling regions. No Significant Improvement with DA.

# Analysis of 1990-91 DA Experiment III




Area-weighted Mean Absolute Error of the monthly-averaged SST anomaly for Jan 1990 through April 1991. In effect, no net reduction in error is seen with respect to the control.

# Improvements to the DA System

- Enhance spread ($S$) based on background variability ($\sigma$): $S \rightarrow (1-c)S + c\sigma$
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix: Boost under-estimated, ensemble-based, flow-dependent covariance with an *a priori*, background estimate.
- Simple bias correction (not yet analyzed)
- Significant improvements in performance of DA system
  - Larger number of obs. contained
  - Significant reduction in SST errors

# Improvements to the DA System

- Enhance spread ($S$) based on background variability ($\sigma$): $S \rightarrow (1-c)S + c\sigma$
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix: Boost under-estimated, ensemble-based, flow-dependent covariance with an *a priori*, background estimate.
- Simple bias correction (not yet analyzed)
- Significant improvements in performance of DA system
  - Larger number of obs. contained
  - Significant reduction in SST errors

# Improvements to the DA System

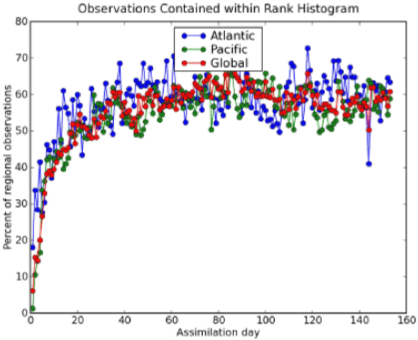
- Enhance spread ($S$) based on background variability ($\sigma$): $S \rightarrow (1 - c)S + c\sigma$
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix: Boost under-estimated, ensemble-based, flow-dependent covariance with an *a priori*, background estimate.
- Simple bias correction (not yet analyzed)
- Significant improvements in performance of DA system
  - Larger number of obs. contained
  - Significant reduction in SST errors

# Improvements to the DA System

- Enhance spread ($S$) based on background variability ($\sigma$): $S \rightarrow (1-c)S + c\sigma$
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix: Boost under-estimated, ensemble-based, flow-dependent covariance with an *a priori*, background estimate.
- Simple bias correction (not yet analyzed)
- Significant improvements in performance of DA system
  - Larger number of obs. contained
  - Significant reduction in SST errors

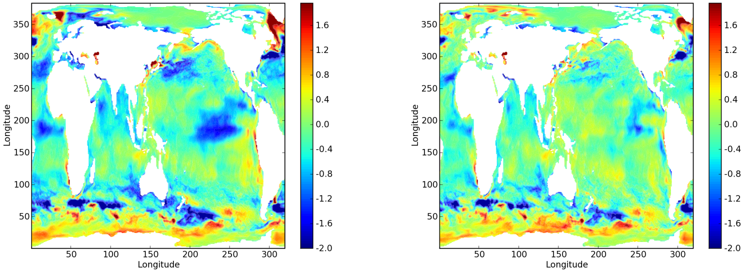
# Improvements to the DA System

- Enhance spread ($S$) based on background variability ($\sigma$): $S \rightarrow (1 - c)S + c\sigma$
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix: Boost under-estimated, ensemble-based, flow-dependent covariance with an *a priori*, background estimate.
- Simple bias correction (not yet analyzed)
- Significant improvements in performance of DA system
  - Larger number of obs. contained
  - Significant reduction in SST errors



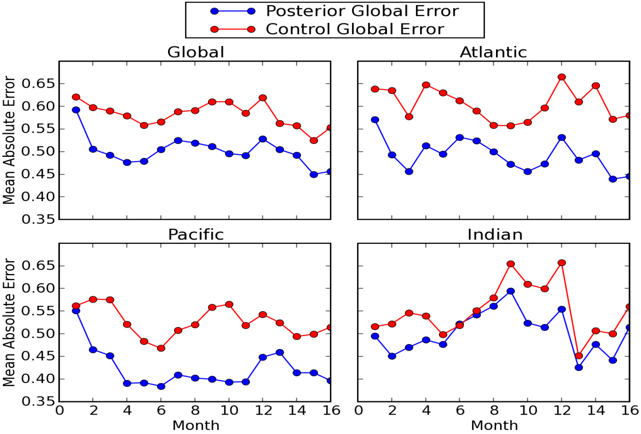

About 60% obs. used compared to $<$ 10% previously

# Improved Assimilation I

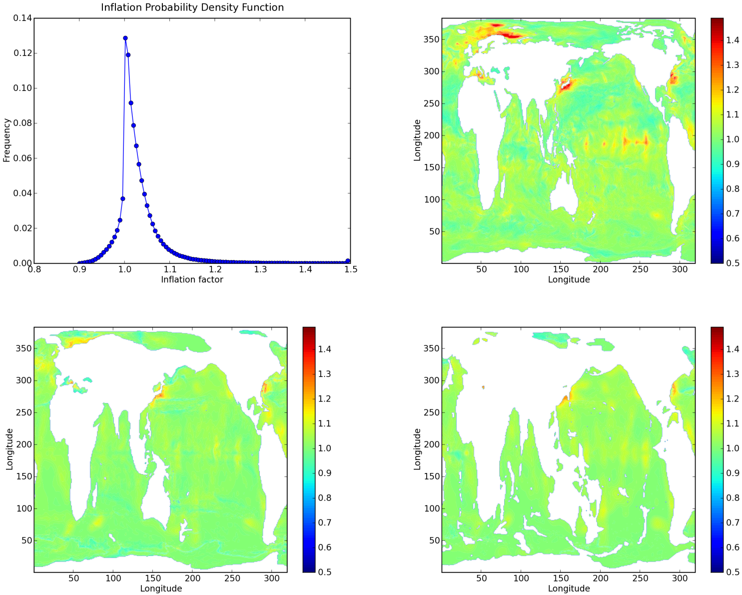



Time-avgd. difference wrt NOAA OI SST V2 **Left**: Same Control Ensemble Run; **Right**: Assimilation Run with Improvements. Significant Improvement over Control Ensemble.

# Improved Assimilation II

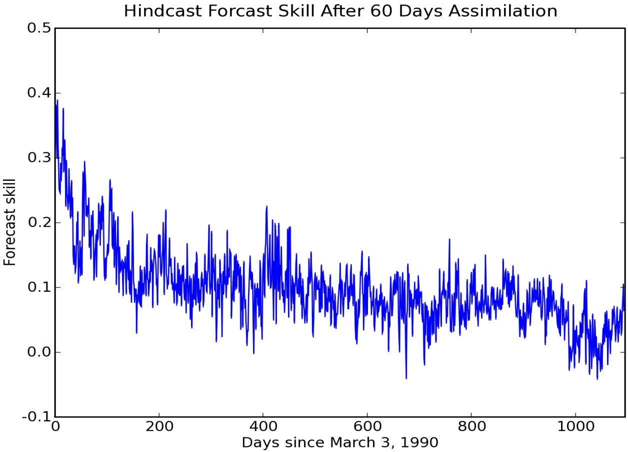



Area-weighted Mean Absolute Error of the monthly-averaged SST anomaly for Jan 1990 through April 1991. Significant reduction in error with respect to control ensemble.




Avg. inflation factors for pot. temp. at surface (top right) and depths of 1000 m (bottom left) and 3200 m

# Hindcast




Faster loss in skill over first six months, followed by a slower decay

# Conclusions

- Modifying spread to include a small fraction of background variability can be useful to improving ensemble diversity
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix.
- While the most serious disadvantage may seem to be a deleterious effect on the spread-skill relation, hindcasts using the ensemble-mean assimilated state shows significant skill

We were able to successfully assimilate WOD observations using an ensemble filter and only a small ensemble. Improved initialization shows skill

# Conclusions

- Modifying spread to include a small fraction of background variability can be useful to improving ensemble diversity
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix.
- While the most serious disadvantage may seem to be a deleterious effect on the spread-skill relation, hindcasts using the ensemble-mean assimilated state shows significant skill

We were able to successfully assimilate WOD observations using an ensemble filter and only a small ensemble. Improved initialization shows skill

# Conclusions

- Modifying spread to include a small fraction of background variability can be useful to improving ensemble diversity
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix.
- While the most serious disadvantage may seem to be a deleterious effect on the spread-skill relation, hindcasts using the ensemble-mean assimilated state shows significant skill

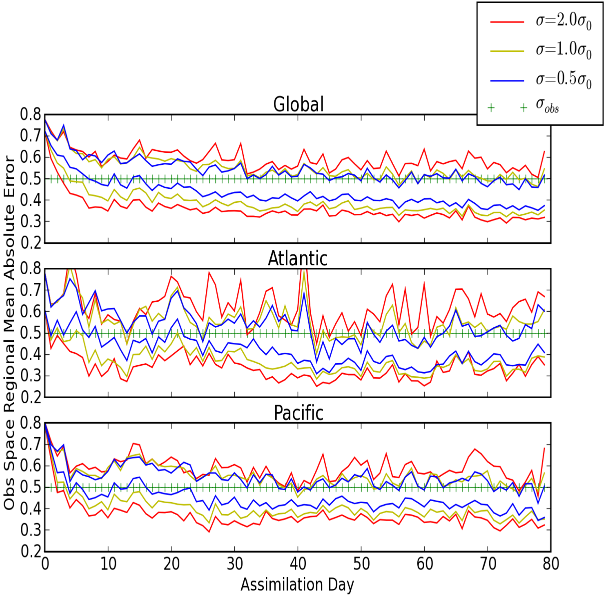
> We were able to successfully assimilate WOD observations using an ensemble filter and only a small ensemble.
> Improved initialization shows skill

# Conclusions

- Modifying spread to include a small fraction of background variability can be useful to improving ensemble diversity
  - Analogous to hybrid methods to boost rank of the forecast error covariance matrix.
- While the most serious disadvantage may seem to be a deleterious effect on the spread-skill relation, hindcasts using the ensemble-mean assimilated state shows significant skill

We were able to successfully assimilate WOD observations using an ensemble filter and only a small ensemble.
Improved initialization shows skill

# Dependence on Amplitude of Specified Background Variability



Anderson, J. (2009).
Spatially and temporally varying adaptive covariance inflation for ensemble filters.
*Tellus A*, 61:72–83.

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A. (2009).
The Data Assimilation Research Testbed: A community facility.
*Bull. Amer. Meteor. Soc.*, pages 1283–1296.

Griffies, S. M., Biastoch, A., Böning, C., Bryan, F., Danabasoglu, G., Chassignet, E. P., England, M. H., Gerdes, R., Haak, H., Hallberg, R. W., Hazeleger, W., Jungclaus, J., Large, W. G., Madec, G., Pirani, A., Samuels, B. L., Scheinert, M., Gupta, A. S., Severijns, C. A., Simmons, H. L., Treguier, A. M., Winton, M., Yeager, S., and Yin, J. (2009).
Coordinated Ocean-ice Reference Experiments (COREs).
*Ocean Modelling*, 26:1–46.

Large, W. G. and Yeager, S. (2009).
The global climatology of an interannually varying air-sea ux data set.
*Climate Dynamics*, 33:341–364.